



Education of the second

NATIONAL BUREAU OF STANDARDS MICROCOPY RESOLUTION TEST CHART



AD-A162 837

SOME CONVERGENCE RESULTS FOR KERNEL-TYPE QUANTILE ESTIMATORS UNDER CENSORING\*

bу

Y. L. Lio and W. J. Padgett

Department of Statistics
University of South Carolina
Columbia, SC 29208

University of South Carolina
Statistics Technical Report No. 108
62G05-14

November, 1985



Research supported by the U.S. Air Force Office of Scientific Research under grant number AFOSR 84-0156 and the U.S. Army Research Office under grant number MIPR ARO 139-85.

ABSTRACT

Based on right-censored data from a lifetime distribution  $F_0$ , a kernel-type estimator of the quantile function  $Q^0(p)$ , defined by  $Q_n(p) = h_n^{-1} \int_0^1 Q_n(t) K((t-p)/h_n) dt$ , is studied. This estimator is smoother than the product-limit quantile function  $Q_n(p) = \inf\{t: \hat{F}_n(t) \geq p\}$ , where  $\hat{F}_n$  denotes the product-limit estimator of  $F_0$  from the censored sample. Under the random censorship model and general conditions on  $h_n$ , K, and  $F_0$ , asymptotic normality of  $Q_n(p)$  and a simpler approximation to it,  $Q_n^*(p)$ , is shown, and mean square convergence of  $Q_n$  is proven. Also, the asymptotic mean equivalence of  $Q_n$  and  $Q_n^*$  is shown.

Accession For	The Irement also cores:
NTIS GRA&I PTIC TAB	asymptote normality, asymptote
d ton	mean I guerolence; and mean
	OUALITY INSPECTED Probability distribution; leywords
or.	Ernel functione
A-I	VIR Bonan und

nief, Tucha para de la comica del la comica de la comica de la comica del la

#### 1. INTRODUCTION AND PRELIMINARIES

For any probability distribution G, denote the quantile function by  $Q(p) = G^{-1}(p) = \inf\{x: G(x) \ge p\}, \ 0 \le p \le 1$ . From a random (uncensored) sample from G, the sample quantile function  $G_n^{-1}(p) = \inf\{x: G_n(x) \ge p\}, \ 0 \le p \le 1$ , has been used to estimate Q(p), where  $G_n$  denotes the sample distribution function. Csörgő (1983) gave many of the known results concerning  $G_n^{-1}(p)$ . Also, Falk (1984) studied the relative deficiency of the sample quantile with respect to kernel-type estimators, and Falk (1985) obtained asymptotic normality for kernel estimators. Yang (1985) obtained some convergence properties of kernel estimators of Q(p) and gave simulation results comparing kernel-type estimators with other estimators. For arbitrarily right-censored data, Sander (1975) proposed estimation of Q(p) by the quantile function of the product-limit estimator, and she and Cheng (1984) derived asymptotic properties while Csörgő (1983) presented strong approximation results for that estimator.

For randomly right-censored data, Padgett (1985) proposed a smooth non-parametric estimator of the quantile function, defined by  $Q_n(p) = h_n^{-1} \int_0^1 \hat{Q}_n(t) K((t-p)/h_n) dt, \text{ where } \hat{Q}_n \text{ denotes the product-limit quantile function and K is an appropriate kernel function. An approximation to <math display="block">Q_n(p), \text{ denoted by } Q_n^*(p), \text{ which is somewhat easier to compute was also studied. The estimator } Q_n, \text{ mentioned briefly by Parzen (1979), was shown to be strongly consistent, and } Q_n \text{ and } Q_n^* \text{ were shown to be almost surely asymptotically equivalent. Lio, Padgett, and Yu (1985) obtained an asymptotic normality result for <math>Q_n$  and showed that  $Q_n$  and  $Q_n^*$  are asymptotically uniformly mean square equivalent under certain conditions.

In this paper, some further asymptotic normality results for  $\mathbf{Q}_n$  and  $\mathbf{Q}_n^\star$  will be given. Also, their asymptotic mean equivalence and the mean

square convergence of  $Q_n$  will be shown. To define these estimators, let  $X_1^0,\dots,X_n^0$  denote the true survival times of n items or individuals that are censored on the right by a sequence  $U_1,U_2,\dots,U_n$ , which in general may be either constants or random variables. It is assumed that the  $X_1^0$ 's are nonnegative independent identically distributed random variables with common unknown distribution function  $F_0$  and unknown quantile function  $Q^0(p) = \xi_p^0$ . The observed right-censored data are denoted by the pairs  $(X_1, \Delta_1)$ ,  $i=1,\dots,n$ , where

$$X_{i} = \min\{X_{i}^{0}, U_{i}^{0}\}, \qquad \Delta_{i} = \begin{cases} 1 & \text{if } X_{i}^{0} \leq U_{i} \\ 0 & \text{if } X_{i}^{0} > U_{i} \end{cases}.$$

Let  $(Z_i, A_i)$ , i=1,...,n, denote the ordered  $X_i$ 's along with their corresponding  $\Delta_i$ 's. A popular estimator of the survival function  $S_0 = 1-F_0$  is the product-limit estimator of Kaplan and Meier (1958), shown to be "self-consistent" by Efron (1967) and defined by

$$\hat{P}_{n}(t) = \begin{cases} 1, & 0 \le t \le Z_{1}, \\ \mathbb{I} & (\frac{n-i}{n-i+1})^{A_{i}}, & Z_{k-1} < t \le Z_{k}, & k=2,...,n \\ 0, & t > Z_{n} \end{cases}$$

Denote the product-limit estimator of  $F_0(t)$  by  $\hat{F}_n(t) = 1 - \hat{P}_n(t)$ , and let  $s_i$  denote the jump of  $\hat{P}_n$  at  $Z_i$ , that is,

$$s_{j} = \begin{cases} 1 - \hat{P}_{n}(Z_{2}), & j = 1 \\ \hat{P}_{n}(Z_{j}) - \hat{P}_{n}(Z_{j+1}), & j = 2, ..., n-1 \\ \hat{P}_{n}(Z_{n}), & j = n. \end{cases}$$

Note that  $s_j=0$  if and only if  $A_j=0$ , j < n, i.e. whenever  $Z_j$  is a censored observation. Also, denote  $S_i=F_n(Z_{i+1})=\sum\limits_{j=1}^{i}s_j$ ,  $i=1,\ldots,n$ , with  $S_0=0$ ,  $Z_0=0$ , and  $Z_{n+1}=Z_n+\epsilon$ , for some positive constant  $\epsilon$ .

It is natural to estimate  $\xi_p^0$  by the product-limit quantile function  $\hat{Q}_n(p) = \hat{\xi}_p^0 = \inf\{t: \hat{F}_n(t) \ge p\}$ . Then the kernel-type estimator  $Q_n(p)$  studied by Padgett (1985) is written as

$$Q_{n}(p) = h_{n}^{-1} \int_{0}^{1} \tilde{Q}_{n}(t)K((t-p)/h_{n})dt$$

$$= h_{n}^{-1} \sum_{i=1}^{n} Z_{i} \int_{s_{i-1}}^{s_{i}} K((t-p)/h_{n})dt, \qquad (1.1)$$

for kernel function K and bandwidth sequence  $\{h_n\}$ . Also, the simpler kernel-type estimator  $Q_n^*$  which is an approximation to (1.1) is defined by  $Q_n^*(p) = h_n^{-1} \sum_{i \in I} Z_i s_i K((S_i - p)/h_n). \tag{1.2}$ 

For the results here, the random right-censorship model will be assumed; that is,  $U_1, \ldots, U_n$  constitute a random sample from a distribution H (usually unknown) and are independent of  $X_1^0, \ldots, X_n^0$ . The distribution function of each  $X_i$ ,  $i=1,\ldots,n$ , is then  $F=1-(1-F_0)(1-H)$ . In addition, some or all of the following conditions will be assumed for the kernel function, bandwidth sequence, and lifetime and censoring distributions:

- (h.1)  $h_n \rightarrow 0$  as  $n \rightarrow \infty$ ;
- (K.1) K(x) is a bounded probability density function which has finite support, i.e. K(x) = 0 for |x| > c for some c > 0;
- (K.2) K is symmetric about zero;
- (K.3) K satisfies a Lipschitz condition, i.e. there exists a constant  $\Gamma$  such that for all x,y,

$$|K(x) - K(y)| \le \Gamma|x - y|;$$

(F.1)  $F_0$  is continuous with density function  $f_0$ . These conditions are not prohibitive and the conditions on  $F_0$  are similar to conditions required by Cheng (1984).

#### 2. ASYMPTOTIC NORMALITY

Lio, Padgett, and Yu (1985) showed that under conditions (h.1),

(K.1)-(K.2), and (F.1), if the derivative  $f_0'$  is continuous at  $\xi_p^0$ ,  $f_0(\xi_p^0) > 0$ , and  $h_n \to 0$ , then for  $0 , where <math>T < \min\{1, T\}$ 

with T = sup {t:  $H(F_o^{-1}(t))<1$ },  $n^{\frac{1}{2}}[Q_n(p)-Q^0(p)] \rightarrow Z$  in distribution as  $n \rightarrow \infty$ , where Z is a normally distributed random variable with mean zero and variance  $\sigma_p^2 = (1-p)^2 \int_0^{\xi_p^0} [1-F(u)]^{-2} dF_o^{\frac{1}{2}}(u)/f_o^2(\xi_p^0)$ . Here  $F_o^{\frac{1}{2}}(u) = \frac{\xi_p^0}{2}$ 

 $P(X_i \le u, \Delta_i = 1)$  is the subdistribution function of the uncensored observations. The condition  $n^{u_i}h_n \to 0$  can be replaced by  $n^{u_i}h_n \to 0$  by using a slightly different proof than that of Lio, Padgett, and Yu (1985).

Define  $Q^O(p,h_n) = h_n^{-1} \int_0^1 Q^O(t) K((t-p)/h_n) dt$  for  $0 \le p \le 1$ . An asymptotic normality result for (1.1) can be obtained without the condition on the rate of convergence of  $h_n$  to zero by centering with  $Q^O(p,h_n)$  instead of  $Q^O(p)$ . This type of centering seems to be required for asymptotic normality of  $Q_n^{*}(p)$ .

THEOREM 2.1. Suppose (h.1), (K.1), (K.2), and (F.1) hold and  $f_0'$  is continuous at  $\xi_p^0$  with  $f_0(\xi_p^0) > 0$ . Then for  $0 , <math display="block">n^{\frac{1}{2}}[Q_n(p) - Q^0(p,h_n)] \rightarrow Z \text{ in distribution as } n \rightarrow \infty, \text{ where } Z \text{ is normally distributed with mean zero and variance } \sigma_p^2.$ 

The proof of Theorem 2.1 follows from the following lemma proven by Lio, Padgett, and Yu (1985) and from Corollary 1 of Cheng (1984).

LEMMA 2.1. Under the conditions of Theorem 2.1

$$\left| \int_0^1 (q_n(t) - q_n(p)) h_n^{-1} K((t-p)/h_n) dt \right| \rightarrow 0 \text{ in probability}$$

as  $n \to \infty$ , where  $q_n(t) = n^{\frac{1}{n}}[\hat{Q}_n(t) - Q^0(t)]$  denotes the product-limit quantile process.

The asymptotic normality of (1.2) follows from Theorem 2.1 and the next

lemma.

LEMMA 2.2. In addition to the conditions of Theorem 2.1, suppose (K.3) holds and  $E(X^{02}) < \infty$ . Then as  $n \to \infty$ ,

$$\sup_{0 \le p \le T} n[Q_n^*(p) - Q_n(p)]^2 \to 0 \text{ in probability,}$$

provided  $h_n^{-4}\sqrt{\log \log n/n} \to 0$ .

PROOF. For  $0 \le p \le T$ , when  $s_i > 0$ , i.e. when  $Z_i$  is uncensored, let  $S_i^*$  be an interior point of the interval  $(S_{i-1}, S_i)$  with probability one so that

$$s_i K \left( \frac{S_i^{\star} - p}{h_n} \right) = \int_{S_{i-1}}^{S_i} K \left( \frac{t-p}{h_n} \right) dt$$
 a.s.

Let  $I_A$  be the indicator function of the set A and let  $i^*$  be the smallest  $i \le n$  such that  $S_{i+1}^{-T} > h_n c$ , where c is the constant in (K.1). If no such i exists, then let  $i^* = n$ . By (K.3),

$$\begin{split} & n[Q_{n}^{\star}(p) - Q_{n}(p)]^{2} I_{[0,T]}(p) \\ &= nh_{n}^{-2} \left\{ \begin{array}{l} n \\ \Sigma \\ i=1 \end{array} Z_{i}s_{i} \left[ K \left( \frac{S_{i}-p}{h_{n}} \right) - K \left( \frac{S_{i}^{\star}-p}{h_{n}} \right) \right] I_{[0,T]}(p) I_{[S_{i}^{\star}-ch_{n},1]}(p) \right\}^{2} \\ &\leq \Gamma^{2} nh_{n}^{-2} \left\{ \begin{array}{l} n \\ \Sigma \\ i=1 \end{array} Z_{i}s_{i}h_{n}^{-1} \left| S_{i}-S_{i}^{\star} \right| I_{[0,T]}(p) I_{[S_{i}^{\star}-ch_{n},1]}(p) \right\}^{2} \\ &\leq \Gamma^{2} h_{n}^{-4} \left[ \sum_{i=1}^{n} Z_{i}^{2}s_{i}^{3} n I_{[0,i^{\star}]}(i) \right] \\ &\leq \Gamma^{2} h_{n}^{-4} \left[ \sum_{i=1}^{n} Z_{i}^{2}s_{i}^{3} n I_{[0,i^{\star}]}(i) \right] \\ &\leq \Gamma^{2} h_{n}^{-4} \left[ \sum_{i=1}^{n} Z_{i}^{2}s_{i}^{3} n I_{[0,i^{\star}]}(i) \right] \end{split}$$

where  $T_{F_0} = \sup \{t: F_0(t) < 1\}$ . From Sander (1975), for  $i \le i^*$ ,  $0 \le ns_i \le [1-H(F_0^{-1}(T))]^{-1} + o_p(1)$ , where  $o_p(1)$  denotes a term converging to zero in probability as  $n \to \infty$ . By the results of Földes and Rejtö (1981),  $\sup_{0 \le x \le T_{F_0}} \left| \hat{F}_n(x) - F_0(x) \right| =$ 

O((log log n/n)<sup>1/2</sup>) with probability one, and by a proof similar to that for Theorem 4.1 of Mauro (1985),  $\sum_{i=1}^{n} Z_{i}^{2} d\hat{F}_{n}(Z_{i}) \rightarrow E(X^{0})$  in probability. Therefore,

 $n[Q_{n}^{*}(p)-Q_{n}(p)]^{2}I_{[0,T]}(p) = O_{p}(h_{n}^{-4}(\log\log n/n)^{\frac{1}{2}}),$  completing the proof.

The following result for asymptotic normality of  $\mathbf{Q}_{n}^{\star}$  is therefore obtained.

THEOREM 2.2. Under the same conditions as in Lemma 2.2, for 0 n^{\frac{1}{2}}[Q\_{n}^{\*}(p)-Q^{0}(p,h\_{n})] \rightarrow Z \text{ in distribution} as n  $\rightarrow \infty$ , where Z is a normally distributed random variable with mean zero and variance  $\sigma_{p}^{2}$ .

#### 3. Asymptotic mean equivalence of $q_n$ and $q_n^*$

It is shown in this section that  $Q_n$  and  $Q_n^*$  are equivalent in the mean and mean squared convergence sense. First, it is proven that  $Q_n^*(p) - Q_n(p)$  converges to zero in the mean uniformly in p for certain choices of  $h_n$ .

For any distribution function G, define  $T_G = \sup\{t: G(t) < 1\}$ .

THEOREM 3.1. Assume that (h.1), (K.1) and (K.3) hold, H and  $F_o$  are continuous with  $E|X^O| < \infty$ , and  $T_{F_o} \le T_H \le \infty$ . Let  $\phi$  be such a function on  $\{0,1-F(T^*)\}$  that  $\phi(x) \ge x$ ,  $\phi(x) \to \phi(0) = 0$  as  $x \to 0^+$ , and  $1 - F_o(t) \le \phi(1 - F(t))$  for  $t \in (T^*, T_{F_o})$ , where  $T^*$  is arbitrary with  $T^* < T_{F_o}$ . Then  $E[|Q_n^*(p) - Q_n(p)|] = O(\phi(d(\frac{\log \log n}{2n})^{\frac{1}{2}})h_n^{-2}),$ 

where d > 1 is some constant.

PROOF. As in the proof of Lemma 2.2, using condition (K.3), with probability one

$$|Q_{n}^{*}(p) - Q_{n}(p)| \leq h_{n}^{-1} \sum_{i=1}^{n} Z_{i} s_{i} |K(\frac{S_{i}^{-p}}{h_{n}}) - K(\frac{S_{i}^{*-p}}{h_{n}})|$$

$$\leq h_{n}^{-2} \Gamma \sum_{i=1}^{n} Z_{i} s_{i}^{2}. \tag{3.1}$$

By continuity of  $F_0$  and using the definitions of  $S_i$  and  $s_i$ , (3.1) is less than or equal to

$$\Pi_{n}^{-2} \int_{0}^{T_{F_{0}}} x |\hat{F}_{n}(x) - \hat{F}_{n}(x^{+})| d\hat{F}_{n}(x) \\
\leq 2\Pi_{n}^{-2} \int_{0}^{T_{F_{0}}} x d\hat{F}_{n}(x) \sup_{0 \leq t \leq T_{F_{0}}} |\hat{F}_{n}(t) - F_{0}(t)|.$$

Thus, by Corollary 2(v) of Csörgö and Horváth (1983) and Theorem 3.1 of Mauro (1985), the conclusion of the theorem follows.

Under similar conditions to those of Theorem 3.1, the asymptotic mean square equivalence of  $\mathbf{Q}_n$  and  $\mathbf{Q}_n^{\star}$  can be obtained for some useful choices of the bandwidth sequence  $\{h_n\}$ .

THEOREM 3.2. Suppose that the conditions of Theorem 3.1 hold, replacing  $\mathbb{E} |\mathbb{X}^0| < \infty$  by  $\mathbb{E}(\mathbb{X}^{02}) < \infty$ . Then

$$E[(Q_n^*(p) - Q_n(p))^2] = 0[\phi^2(d(\frac{\log \log n}{n})^{\frac{1}{2}}h_n^{-4}].$$

PROOF. By an argument similar to the proof of Theorem 3.1, with probability one

$$(Q_n^{\star}(p) - Q_n(p))^2 \le 4\Gamma^2 h_n^{-4} \int_0^{T_p} o_x^2 d\hat{F}_n(x) \sup_{0 \le t \le T_p} |\hat{F}_n(t) - F_o(t)|^2.$$

Again, the conclusion follows from Corollary 2(v) of Csörgö and Horváth (1983) and Theorem 3.1 of Mauro (1985).

From Theorems 3.1 and 3.2, if  $\phi(d(\log\log n/2n)^{\frac{1}{2}})h_n^{-2} \to 0$  as  $n \to \infty$ , then  $Q_n^*(p)$  and  $Q_n(p)$  are asymptoically equivalent in the mean. If  $\phi$  is chosen so that  $\phi(x) = (x/k)^{1/(1+\gamma)}$ , for some  $0 < k \le 1$  and  $\gamma \ge 0$ , as in example (1) of Csörgö and Horváth (1983, p. 416), then (for  $\gamma=0$  and k=1) the condition above becomes (log log n/n)  $h_n^{\frac{1}{2}}h_n^{-2} \to 0$ , which is satisfied by  $h_n \approx Dn^{-b}$  for 0 < b < 4 and some positive constant D, for example.

#### 4. MEAN SQUARE COVERGENCE

The following theorem yields the mean square convergence of  $Q_n(p)$  to  $Q^0(p)$  for appropriate choices of  $h_n$ . Also, combining Theorem 3.2 with Theorem 4.1 below gives conditions under which  $Q_n^{\star}(p)$  converges in mean square to  $Q^0(p)$ .

THEOREM 4.1. Let  $p_0$  be such that  $0 \le p_0 < \min \{1, T_0^{-1}\}$ . Suppose (h.1)

and (K.1) hold,  $F_0$  is differentiable on some neighborhood of  $\xi_p^0$ ,  $f_0$  is continuous at  $\xi_p^0$  with  $f_0(\xi_p^0) > 0$ , and  $E(X^{04}) < \infty$ . Then for  $0 \le p < p_0$ ,  $E\{[Q_n(p) - Q^0(p)]^2\} \le g(n,h_n), \text{ where } g(n,h_n) = 0(h_n^2) + 0(h_n^{\frac{1}{2}}n^{-5/4}) + 0(h_n^{-5/4}(\log n)^{3/4}) + 0(n^{-5/6}(\log n)^{5/6}) + 0(n^{-\frac{1}{4}}(c_1(1-p)^2 + c_2(1-p)h_n + c_3h_n^2)) \text{ for some positive constants } c_1, c_2, \text{ and } c_3.$ 

It should be noted that an example of a bandwidth sequence  $\{h_n\}$  which will give  $g(n,h_n) \to 0$  as  $n \to \infty$  is  $h_n \approx c_n n^{-\delta}$ ,  $0 < \delta < 5/2$ , where  $c_n > 0$  is bounded by some positive constant d.

The proof of Theorem 4.1 is obtained from the following three lemmas.

LEMMA 4.1. Suppose (K.1) holds,  $F_o$  is differentiable on some neighborhood of  $\xi_p^o$  with  $f_o(\xi_p^o)>0$ . Then

$$\{\int_{-c}^{c} [Q^{o}(p+h_{n}u) - Q^{o}(p)] K(u)du\}^{2} = O(h_{n}^{2}).$$

PROOF. By the conditions on  $F_0$  and  $f_0$ , there exists a neighborhood A(p) of  $\xi_p^0$  so that  $A = \sup_{t \in A(p)} [f_0(Q^0(t))]^{-1} < \infty$ . Hence, by condition (K.1), the conclusion of the lemma follows.

LEMMA 4.2. Under the conditions of Theorem 4.1,

$$\begin{split} \big| \mathbb{E} \big\{ \int_{-c}^{c} \big[ \hat{Q}_{n}(p + h_{n}u) - Q^{0}(p + h_{n}u) \big] K(u) du \big| \\ & \leq O(h_{n}^{-\frac{1}{2}} n^{-5/4}) + O((\log n/n)^{3/4}). \end{split}$$

PROOF. Let  $U_n$  denote the product-limit (PL) quantile process (Csorgö, 1983, Eq. (8.1.18), p. 118), and for each n choose  $\varepsilon_n = (\log n/n)^{\frac{1}{12}}$ . Define the events  $A_n = [\sup_{-c \le u \le c} |U_n(p+h_n u) - (p+h_n u)| > \varepsilon_n]$ . By Cheng (1984), for

large n, 
$$P[A_n] = O(n^{-5/2})$$
.

$$\begin{aligned} & \text{Write } \mathbb{E}\{\int_{-c}^{c} [\hat{Q}_{n}(p+h_{n}u) - Q^{0}(p+h_{n}u)]K(u)du\} = \mathbb{E}_{1} + \mathbb{E}_{2}, \text{ where } \\ & \mathbb{E}_{1} = \mathbb{E}\{\int_{-c}^{c} [Q^{0}(U_{n}(p+h_{n}u) - Q^{0}(p+h_{n}u)]K(u)du \cdot \mathbb{I}_{A_{n}}\} \end{aligned}$$

and

$$E_2 = E\{\int_{-c}^{c} [Q^o(U_n(p+h_n u)) - Q^o(p+h_n u)]K(u)du \cdot I_{A_n}^c\}.$$

Using the Holder inequality,

$$|\mathbb{E}_{1}| \leq \{\mathbb{E}[\int_{-c}^{c} Q^{0}(U_{n}(p+h_{n}u))K(u)du]^{2} \mathbb{P}[\mathbb{A}_{n}]\}^{\frac{1}{2}}$$

$$\leq \{\sup_{u} K(u) n^{-5/2} \mathbb{E}[h_{n}^{-1}\int_{-c}^{c} \hat{Q}_{n}^{2}(t)dt]\}^{\frac{1}{2}}.$$

Thus, by Theorem 3.1 of Mauro (1985),  $|E_1| = O(h_n^{-\frac{1}{2}}n^{-5/4})$ .

Now, for E2, using Taylor's expansion, there exists a  $\xi$  between  $\textbf{U}_n(\textbf{p}+\textbf{h}_n\textbf{u})$  and  $\textbf{p}+\textbf{h}_n\textbf{u}$  such that

+ 
$$n^{-\frac{1}{2}} \tilde{\alpha}_{n}^{(p+h_{n}u)} [K(u)du \cdot I_{A_{n}^{c}}]$$
  
-  $\int_{-c}^{c} [f_{0}(Q^{0}(\xi))]^{-1} n^{-\frac{1}{2}} \tilde{\alpha}_{n}^{(p+h_{n}u)} K(u)du I_{A_{n}^{c}}],$ 

where  $\alpha_n^-$  denotes the uniform PL - empirical process (Csorg5, 1983, p. 117). For n sufficiently large,  $\left|1/f_0(Q^0(\xi))\right| \le A$ , where A is defined in the proof of Lemma 4.1, and by Cheng's (1984) Theorem 2, for large n,

$$\sup_{-c \le u \le c} [U_n(p+h_n u) - (p+h_n u) + n^{-\frac{1}{2}\alpha} (p+h_n u)]$$

$$= 0((\log n/n)^{3/4}). \tag{4.1}$$

Next.

$$|E\int_{-c}^{c} n^{-\frac{1}{2}\alpha} \alpha_{n}(p+h_{n}u)K(u)du/f_{0}(Q^{0}(\xi))I_{A_{n}^{c}}|$$

$$\leq n^{-\frac{1}{2}} A \mathbb{E} \{ \int_{-c}^{c} |\tilde{\alpha}_{n}(p+h_{n}u) - n^{-\frac{1}{2}} K^{*}(p+h_{n}u,n) | K(u) du \} + n^{-\frac{1}{2}} | \mathbb{E} \int_{-c}^{c} n^{-\frac{1}{2}} K^{*}(p+h_{n}u,n) K(u) du \mathbb{I}_{A_{n}^{c}} |,$$
(4.2)

where K\*(t,s) denotes the generalized Kiefer process (Csorgo, 1983,

p. 118). By Theorem 8.1.1 of Csörgö (1983) (or Burke, Csörgö, and Horvath, 1981),

$$\sup_{-c \le u \le c} |\alpha(p+h_n u) - n^{-h} K^*(p+h_n u, n)| \stackrel{a.s.}{=} 0(n^{-1/3} (\log n)^{5/2}). \tag{4.3}$$

Therefore, from (4.1)-(4.3), for large n,

$$\begin{split} |\mathbb{E}_{2}| &\leq \mathbb{E}\{\mathbb{A}\int_{-c}^{c} |[\mathbb{U}_{n}(p+h_{n}u) - (p+h_{n}u) + n^{-\frac{1}{2}}\tilde{\alpha}_{n}(p+h_{n}u)]\mathbb{K}(u)|du \\ &+ n^{-\frac{1}{2}}|\mathbb{E}\int_{-c}^{c} n^{-\frac{1}{2}}\mathbb{K}^{*}(p+h_{n}u,n)\mathbb{K}(u)du\mathbb{I}_{A_{n}^{c}}| \\ &+ 0(n^{-1/3}(\log n)^{5/2})n^{-\frac{1}{2}} \\ &\leq 0((\log n/n)^{3/4}) + 0(n^{-5/6}(\log n)^{5/2} \\ &= 0((\log n/n)^{3/4}), \end{split}$$

since  $\mathbb{E}\int_{-c}^{c} n^{-\frac{1}{2}} |K^*(p+h_n^u,n)| K(u) du < \infty$ , and the proof is completed.

LEMMA 4.3. Suppose the conditions of Theorem 4.1 hold. Then

$$\begin{split} \mathbb{E}\{[\int_{-c}^{c}(\hat{\mathbf{Q}}_{n}(\mathbf{p}+\mathbf{h}_{n}\mathbf{u}) - \mathbf{Q}^{0}(\mathbf{p}+\mathbf{h}_{n}\mathbf{u})K(\mathbf{u})\mathrm{d}\mathbf{u}]^{2}\} \\ &\leq 0(n^{-5/6}(\log n)^{5/2}) + 0(\mathbf{h}_{n}^{-4}\mathbf{h}_{n}^{-5/4}) \\ &+ 0(n^{-4}(\mathbf{c}_{1}(1-\mathbf{p})^{2} + \mathbf{c}_{2}(1-\mathbf{p})\mathbf{h}_{n} + \mathbf{c}_{3}\mathbf{h}_{n}^{2})). \end{split}$$

$$\begin{split} \mathbb{P}ROOF. \quad &\text{As in the proof of Lemma 4.2, write} \\ \mathbb{E}\{[\int_{-c}^{c}(\hat{\mathbf{Q}}_{n}(\mathbf{p}+\mathbf{h}_{n}\mathbf{u}) - \mathbf{Q}^{0}(\mathbf{p}+\mathbf{h}_{n}\mathbf{u}))K(\mathbf{u})\mathrm{d}\mathbf{u}]^{2}\} = \mathbb{E}_{3} + \mathbb{E}_{4}, \end{split}$$

$$\text{where} \\ &\mathbb{E}_{3} = \mathbb{E}\{[\int_{-c}^{c}(\hat{\mathbf{Q}}_{n}(\mathbf{p}+\mathbf{h}_{n}\mathbf{u}) - \mathbf{Q}^{0}(\mathbf{p}+\mathbf{h}_{n}\mathbf{u}))K(\mathbf{u})\mathrm{d}\mathbf{u}]^{2}\mathbf{I}_{A}^{c}\} \\ \text{and} \\ &\mathbb{E}_{4} = \mathbb{E}\{[\int_{-c}^{c}(\hat{\mathbf{Q}}_{n}(\mathbf{p}+\mathbf{h}_{n}\mathbf{u}) - \mathbf{Q}^{0}(\mathbf{p}+\mathbf{h}_{n}\mathbf{u}))K(\mathbf{u})\mathrm{d}\mathbf{u}]^{2}\mathbf{I}_{A}^{c}\} \\ \text{Now,} \\ &|\mathbb{E}_{4}| \leq \mathbb{E}\{[\int_{-c}^{c}(\hat{\mathbf{Q}}_{n}^{2}(\mathbf{p}+\mathbf{h}_{n}\mathbf{u})K(\mathbf{u})\mathrm{d}\mathbf{u})]\mathbf{I}_{A}^{n}\} \\ &\leq \{\mathbb{E}[(\int_{-c}^{c}(\hat{\mathbf{Q}}_{n}^{2}(\mathbf{p}+\mathbf{h}_{n}\mathbf{u})K(\mathbf{u})\mathrm{d}\mathbf{u})]^{\frac{1}{2}}\mathbf{n}^{-5/4} \\ &\leq \{\mathbb{E}[\int_{-c}^{c}(\hat{\mathbf{Q}}_{n}^{2}(\mathbf{p}+\mathbf{h}_{n}\mathbf{u})K(\mathbf{u})\mathrm{d}\mathbf{u}]\}^{\frac{1}{2}}\mathbf{n}^{-5/4} \\ &\leq [\sup_{\mathbf{u}} K(\mathbf{u})]^{\frac{1}{2}}\{\mathbf{h}_{n}^{-1}\mathbb{E}[\hat{\mathbf{Q}}_{n}^{4}(\mathbf{t})\mathrm{d}\mathbf{t}\}^{\frac{1}{2}}\mathbf{n}^{-5/4} \\ &\leq 0(\mathbf{h}_{n}^{-\frac{1}{2}}\mathbf{n}^{-5/4}) \end{split}$$

by Hölder's inequality and Mauro's (1985) Theorem 3.1.

Next, using the Taylor expansion as in the proof of Lemma 4.2,

$$\begin{split} |E_3| &\leq |E\{[\int_{-c}^c (f_0(Q^0(\xi)))^{-1}(U_n(p+h_nu) - (p+h_nu) \\ &+ n^{-\frac{1}{2}}\tilde{\alpha}_n(p+h_nu))K(u)du]^2\}| \\ &+ |E\{[\int_{-c}^c (f_0(Q^0(\xi)))^{-1}n^{-\frac{1}{2}}\tilde{\alpha}_n(p+h_nu)K(u)du]^2\}| \\ &+ 2|E\{[\int_{-c}^c (f_0(Q^0(\xi)))^{-1}[U_n(p+h_nu) - (p+h_nu) \\ &+ n^{-\frac{1}{2}}\tilde{\alpha}_n(p+h_nu)]K(u)du \\ &\times \int_{-c}^c (f_0(Q^0(\xi)))^{-1}n^{-\frac{1}{2}}\tilde{\alpha}_n(p+h_nu)K(u)du\}| \\ &= E_{31} + E_{32} + E_{33}. \end{split}$$
 Now, as in the proof of Lemma 4.2, 
$$|E_{31}| \leq [0(\log n/n)^{3/4})]^2 = 0((\log n/n)^{3/2}). \end{split}$$

Also,

$$\begin{split} E_{32} & \leq n^{-\frac{1}{4}} \left| \mathbb{E} \{ [\int_{-c}^{c} (f_{o}(Q^{o}(\xi)))^{-1} (\tilde{\alpha}_{n}(p + h_{n}u) - n^{-\frac{1}{4}} K^{*}(p + h_{n}u, n)) K(u) du ]^{2} \} \right| \\ & + n^{-\frac{1}{4}} \left| \mathbb{E} \{ \int_{-c}^{c} (f_{o}(Q^{o}(\xi)))^{-1} n^{-\frac{1}{4}} K^{*}(p + h_{n}u) K(u) du \}^{2} \right| \\ & + n^{-\frac{1}{4}} 2 \left| \mathbb{E} \{ \int_{-c}^{c} (f_{o}(Q^{o}(\xi)))^{-1} [\tilde{\alpha}_{n}(p + h_{n}u) - n^{-\frac{1}{4}} K^{*}(p + h_{n}u, n)] K(u) du \right| \\ & - n^{-\frac{1}{4}} K^{*}(p + h_{n}u, n) K(u) du \\ & \leq n^{-\frac{1}{4}} [0(n^{-1/3}(\log n)^{5/2})]^{2} \\ & + A^{2} n^{-\frac{1}{4}} \mathbb{E} \{ \int_{-c}^{c} n^{-1} K^{*2}(p + h_{n}u, n) K^{2}(u) du \} \\ & + 2A^{2} n^{-\frac{1}{4}} 0(n^{-1/3}(\log n)^{5/2}) \mathbb{E} \int_{-c}^{c} n^{-\frac{1}{4}} [K^{*}(p + h_{n}u, n)] K(u) du. \end{split} \tag{4.5}$$

Since  $n^{-1/2}K^*(t,n)$  is a Gaussian process in t, the last expectation in (4.5) is finite, and by Fubini's theorem and section 8.2 of Csörgö (1983), for n sufficiently large,

$$E\{\int_{-c}^{c} n^{-1} K^{*2}(p+h_{n}u,n)K(u)du\}$$

$$= \int_{-c}^{c} [1-(p+h_{n}u)]^{2} \left( \int_{0}^{p+h_{n}u} \frac{dx}{(1-x)^{2}(1-H(Q^{0}(x)))} \right) K(u)du$$

$$\leq \left[ \int_{0}^{p_{0}} \frac{dx}{(1-x)^{2}(1-H(Q^{0}(x)))} \right] \left[ \int_{-c}^{c} (1-p)^{2} K(u)du + 2(1-p)h_{n} \int_{-c}^{c} uK(u)du + h_{n}^{2} \int_{-c}^{c} u^{2} K(u)du \right]. \tag{4.6}$$

Therefore, from (4.6) and condition (K.1), (4.5) yields for appropriate constants  $c_1$ ,  $c_2$ , and  $c_3$ 

$$\begin{split} |\mathbb{E}_{32}| & \leq 0 (n^{-7/2} (\log n)^{5/2}) + n^{-\frac{1}{4}} [c_1 (1-p)^2 + c_2 (1-p)h_n + c_3 h_n^2]. \\ \text{Similarly, since } |\mathbb{E}_{J_{-c}}^c (f_0 (Q^0(\xi)))^{-1} n^{-\frac{1}{4} \tilde{\alpha}}_n (p+h_n u) K(u) du| < \infty, \\ |\mathbb{E}_{33}| & \leq 0 (n^{-5/6} (\log n)^{5/2}). \end{split}$$

Therefore, combining these results for  $E_3$ ,

$$\begin{split} \mathbb{E}\{\int_{-c}^{c} (\hat{Q}_{n}(p+h_{n}u) - Q^{0}(p+h_{n}u))K(u)du]^{2}\} \\ &\leq 0(h_{n}^{-\frac{1}{2}}n^{-5/4}) + 0((\log n/n)^{3/2}) + 0(n^{-7/2}(\log n)^{-5/2}) \\ &+ 0(n^{-\frac{1}{4}}[c_{1}(1-p)^{2} + c_{2}(1-p)h_{n} + c_{3}h_{n}^{2}]) + 0(n^{-5/6}(\log n)^{5/2}) \end{split}$$

$$= 0(h_n^{-\frac{1}{6}}n^{-5/4}) + 0(n^{-5/6}(\log n)^{5/2}) + 0(n^{-\frac{1}{6}}[c_1(1-p)^2 + c_2(1-p)h_n + c_3h_n^2]),$$
 completing the proof.

The proof of Theorem 4.1 is then obtained from Lemmas 4.1-4.3 by writing, for sufficiently large n,

$$\begin{split} \mathbb{E}\{\{Q_{n}(p) - Q^{0}(p)\}^{2}\} \\ &= \mathbb{E}\{\{\int_{-c}^{c} (\hat{Q}_{n}(p+h_{n}u) - Q^{0}(p+h_{n}u))K(u)du\}^{2}\} \\ &+ \{\int_{-c}^{c} (Q^{0}(p+h_{n}u) - Q^{0}(p))K(u)du\}^{2} \\ &+ 2\int_{-c}^{c} [Q^{0}(p+h_{n}u) - Q^{0}(p)]K(u)du \\ &\times \mathbb{E}\{\int_{-c}^{c} [\hat{Q}_{n}(p+h_{n}u) - Q^{0}(p+h_{n}u)]K(u)du\}. \end{split}$$

It should be noted that the conditions of Theorem 4.1, as well as those of the other theorems in this paper, are not restrictive and are similar to conditions for results for right-censored data obtained by previous authors. See Chapter 8 of Csörgö (1983) for various references and Cheng (1984), for example.

#### REFERENCES

- BURKE, M.D., CSÖRGÖ, S. and HORVÁTH, L. (1981). Strong Approximations of Some Biometric Estimates under Random Censorship. Zeitschrift fur Wahrscheinlichkeitstheorie und Verwande Gebiete, 56, 87-112.
- CHENG, K.F. (1984). On Almost Sure Representations for Quantiles of the Product Limit Estimator with Applications. Sankhya, Ser. A 46, 426-443.
- CSORGO, M. (1983). Quantile Processes with Statistical Applications. CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, PA.
- CSÖRGÖ, S. and HORVÁTH, L. (1983). The Rate of Strong Uniform Consistency for the Product-Limit Estimator. Zeitschrift für Wahrscheinlichkeitstheorie und Verwande Gebiete, 62, 411-426.
- EFRON, B. (1967). The Two-Sample Problem with Censored Data. Proceedings of the Fifth Berkeley Symposium, 4, 831-853.
- FALK, M. (1984). Relative Deficiency of Kernel Type Estimators of Quantiles. Annals of Statistics, 12, 261-268.
- FALK, M. (1985). Asymptotic Normality of the Kernel Quantile Estimator. Annals of Statistics, 13, 428-433.
- FÖLDES, A., and REJTÖ, L. (1981). A LIL Type Result for the Product Limit Estimator. Zeitschrift für Wahrscheinlichkeitstheorie und Verwande Gebiete, 56, 76-86.
- KAPLAN, B.L., and MEIER, P. (1958). Nonparametric Estimation from Incomplete Observations. <u>Journal of the American Statistical</u> Association, 53, 457-481.
- LIO, Y.L., PADGETT, W.J., and YU, K.F. (1985). On the Asymptotic Properties of a Kernel-Type Quantile Estimator from Censored Samples. <u>Journal of Statistical Planning and Inference</u>, in press.
- MAURO, D. (1985). A Combinatoric Approach to the Kaplan-Meier Estimator. Annals of Statistics, 13, 142-149.
- PADGETT, W.J. (1985). A Kernel-Type Estimator of a Quantile Function from Right-Censored Data. <u>Journal of the American Statistical Association</u>, in press.
- PARZEN, E. (1979). Nonparametric Statistical Data Modeling. <u>Journal of</u>
  the American Statistical Association, 74, 105-121.
- SANDER, J. (1975). The Weak Convergence of Quantiles of the Product Limit Estimator. Technical Report 5, Stanford University, Dept. of Statistics.

YANG, S.S. (1985). A Smooth Nonparametric Estimator of a Quantile Function.

<u>Journal of the American Statistical Association</u>, in press.

### AD-A162837

ECHBITY	CI ACC	FICAT	ION OF	THIS	PAGE

ECURITY CLASSIFICATION OF THIS PAGE										
REPORT DOCUMENTATION PAGE										
1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED			1b. RESTRICTIVE MARKINGS							
2a. SECURIT	TY CLASSIFIC	CATION AUTHORITY		3. DISTRIBUTION/AVAILABILITY OF REPORT						
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE			Approved for public release; distribution unlimited.							
_		ZATION REPORT NUM		5. MONITORING OR						
Stat.	Tech Re	p. No. 108 (626	G05-14)	AFOST (1)						
_			6b. OFFICE SYMBOL (If applicable)	Air Force Office of Scientific Research						
	•	and ZIP Code)		7b. ADDRESS (City, State and ZIP Code)						
University of South Carolina Columbia, SC 29208			Directorate of Mathematical & Information Sciences, Bolling AFB DC 20332							
	F FUNDING/S	SPONSORING	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT I	NSTRUMENT ID	ENTIFICATION NU	MBER			
	ARO		NM	AFOSR-84-0	156					
Sc. ADDRES	SS (City, State	and ZIP Code)		10. SOURCE OF FUN						
Bolli	ng AFB	DC 20332		PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.			
11. TITLE	Include Securit	v Classification)		61102F	2304	A5				
11. TITLE (Include Security Classification)  Some Convergence Results for Kernel-Type Quantile Estimators under Censoring										
	Lio and	(S) W. J. Padgett	•							
13a TYPE OF REPORT 13b. TIME COVERED				14. DATE OF REPORT (Yr., Mo., Dey) 15. PAGE COUNT						
FROM TO 1985 November 15										
17.	COSATI	CODES	18. SUBJECT TERMS (C	ontinue on reverse if ne	cessary and identi	fy by block number				
FIELD	GROUP	SUB. GA.	Random right-c	right-censorship; Kernel estimation; Product-limit						
				antile function; Asymptotic normality; Mean convergence;						
19. ABSTRACT (Continue on reverse if necessary and identify by block number)  Based on right-censored data from a lifetime distribution $F_0$ , a kernel-type estimator of the quantile function $Q^0(p)$ , defined by $Q_n(p) = h_n^{-1} \int_0^1 \hat{Q}_n(t) K((t-p)/h_n) dt$ , is studied.										
			n the product-li		•					
	^	^					ensored			
$\inf\{t\colon F_n(t)\geq p\}$ , where $F_n$ denotes the product-limit estimator of $F_0$ from the censored sample. Under the random censorship model and general conditions on $h_n$ , K, and $F_0$ ,										
asymptotic normality of $Q_n(p)$ and a simpler approximation to it, $Q_n(p)$ , is shown, and mean										
square convergence of $Q_n$ is proven. Also, the asymptotic mean equivalence of $Q_n$ and $Q_n$										
is shown.										
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT			21. ABSTRACT SECURITY CLASSIFICATION							
UNCLASSIFIED/UNLIMITED XX SAME AS RPT.   OTIC USERS			UNCLASSIFIED							
22a. NAME OF RESPONSIBLE INDIVIDUAL			22b. TELEPHONE NI (Include Area Co		22c. OFFICE SYM	BOL				
Maj. Brian W. Woodruff			(202) 767-50		NM					

# END

## FILMED

2-86

DTIC